

A Little Less Interaction, A Little More Action: A Modular Framework for Network Troubleshooting

István Pelle^{1,*}, Felicián Németh¹ and András Gulyás²

¹Budapest University of Technology and Economics, Hungary, HSNLab, Dept. of
Telecommunications and Media Informatics

²Budapest University of Technology and Economics, Hungary, HSNLab, Dept. of
Telecommunications and Media Informatics and MTA-BME Information Systems Research
Group and was supported by the János Bolyai Fellowship of the Hungarian Academy of Sciences

*Corresponding author: István Pelle (pelle@tmit.bme.hu)

March 1, 2017

Abstract

An ideal network troubleshooting system would be an almost fully automated system, monitoring the whole network at once, feeding the results to a knowledge-based decision making system that suggests actions to the operator or corrects the failure automatically. Reality is quite the contrary: operators separated in their offices try to track down complex networking failures in their own way, which is generally a long sequence of manually edited parallel shell commands (mostly ping, traceroute, route, iperf, ofctl etc.). This process requires operators to be “masters of complexity” (which they often are) and continuous interaction. In this paper we aim at narrowing this huge gap between vision and reality by introducing a modular framework capable of (i) formalizing troubleshooting processes as the concatenation of executable functions [called troubleshooting graphs (TSGs)], (ii) executing these graphs via an interpreter, (iii) evaluating and navigating between the outputs of the functions and (iv) sharing troubleshooting know-hows in a formalized manner.

1 Introduction

Troubleshooting a communication network was never an easy problem. Finding causes of errors and failures, tracking down misconfigurations in the increasingly complex interconnection networks of heterogeneous networking devices is quite a challenge. What is more, the prevalence of increasingly complex software components, due to the upcoming software defined networks (SDNs), adds distributed software debugging as an additional issue to deal with. To cope with this increasing complexity, the networking research community suggests the use of knowledge-based decision support together with the standard network monitoring and diagnostic tools, and the conversion of troubleshooting into a highly automated process. Reality seems to reside very far away from this vision. Real operators tend to use the most basic diagnostic tools for monitoring the network, and rely on their own brilliance and programming skills when digging out the root causes of errors in an ad-hoc manner from the reports of these tools. Even if this approach works well in practice, it requires extremely skilled operators who can keep in mind all the details of the network under scrutiny and their continuous interaction usually is wasted on rummaging in the logs of the tools used by them.

As we see, the reason for this huge gap between the ideas and reality is threefold. First, there is no usable, implementation oriented formal description of the troubleshooting processes. Secondly, there is no platform capable of executing formally defined troubleshooting processes while giving prompt and systematic access to the outputs of the used tools. Finally, there is no existing platform that could integrate existing troubleshooting tools and decision support methodologies in a flexible manner. In lack of formalism and integrated execution platform, operators cannot share and re-use each others troubleshooting know-hows in a structured way, thus knowledge is not accumulated but remains sporadic as operators treat every specific failure in their own ad-hoc way.

The purpose of this paper is narrowing the gap between troubleshooting visions and real life solutions. In this respect, our contribution will be threefold. First, we propose a formalization of troubleshooting processes in the form of troubleshooting graphs (TSGs), which let operators describe the steps of tracking down a specific network failure in a structural manner with a very small effort compared to the implementation of it. Once created, TSGs can make their solutions ready-to-share and re-usable. We also define a language for a text-based representation of TSGs. Secondly, we propose a modular execution framework capable of running TSGs and giving on demand fast semantic navigation among the outputs of the tools used in the troubleshooting process. Finally, we present a complete prototype system (called *Epoxide*) capable of defining, executing and analyzing TSGs. Case studies using Epoxide are also given.

The rest of our paper is structured as follows: in Section 2 we give a brief overview on the related work in both literature and practice. Section 3 lists the principles of our proposed modular troubleshooting framework, followed by the illustration of its operation over an everyday example in Section 4. Section 5 presents the fundamentals of our prototype, Epoxide, which is complemented with some illustrative case studies in Section 6. Finally, we conclude the paper and give directions for future works in Section 7.

2 State of the Art in Network Troubleshooting

From the great volume of related literature we highlight here the two main constituents of troubleshooting systems. The first is clearly the area of *network monitoring and diagnostic tools*, of which main purpose is to seek for symptoms of specific failures. The palette is very broad here, ranging from the most basic tools—like ping, traceroute, tcpdump, netstat, nmap [1] or GNU Debugger (GDB)—, through monitoring protocols—such as SNMP and RMON [1]—, configuration files and analyzers—such as Splat [2]—, performance measurement tools—such as iperf [1]—and packet analyzers—like Wireshark—, to the more complex ones—such as NetFlow, HSA [3, 4] and ATPG [5]. On top of these, SDN specific tools have added a whole new segment targeted to investigate specific parts of the architecture. Tools such as Anteater [6], OFRewind [7], NetSight [8], VeriFlow [9], NICE [10], SOFT [11], FORTNOX [12] and OFTEN [13] all fill a niche in SDN troubleshooting.

One level up, the output symptoms of these tools can be aggregated and fed into different *automatic reasoning solutions*. The first representatives of these were created as early as the second half of the 1980s [14] targeting the discovery of failures in telecommunication networks. Early on, rule-based methods were used to resolve issues by using *if-then* statements [15]. Later case-based reasoning [16] and model-based [17] methods were developed. The former utilized a collection of previous cases as a basis for failure analysis, while the latter used models of structural and functional behavior to reason about network issues. Fault-symptom graphs [18] and dependency or causality graphs [19, 20] introduced the concept of tracking failures using graphs that created connections between symptoms, detection and root causes. This concept led to the application of Bayesian networks [21, 22] where belief—in the most probable failure root cause—propagation is based on a probabilistic model.

2.1 What We See in Current Practice

Despite the readily available set of advanced troubleshooting tools and decision support mechanisms, operators seem to use the most rudimentary tools (like ping, traceroute, tcpdump etc.) while they completely rely on their minds as a knowledge-base. For testing this, we conducted an in-house survey querying which type of problems local administrators run into most frequently and what network troubleshooting tools they use most commonly. The results that we found were completely in accordance with those outlined in [23]. Most problems were caused by connectivity issues that arose from a variety of reasons ranging from hardware failures to configuration changes that became necessary due to security issues. Used troubleshooting tools show similarities also: mostly simple task specific tools are utilized, in certain cases combining them in a script to explore typical failures. Network information is usually stored in simple spreadsheets and proprietary monitoring or troubleshooting tools are used only when they have a low cost—or are preferably free. We found that automatic tools are less frequently used and manual troubleshooting dominates problem solving.

To get a deeper sense of the process, think of an operator logging into different devices and running heterogeneous software tools (e.g. the tools listed above) to analyze the problem. Network conditions are monitored simultaneously in multiple shells and after a painful rummaging in the tools’ outputs, new commands are evoked on new devices. We argue that this approach has its own merits and drawbacks. It is extremely flexible as the operator uses the tool of her choosing in the way and logic she sees fit. On the negative side, it is quite ineffective, unorganized and anything but re-usable. The operator’s time is spent mostly on filtering and finding the correlation between the different outputs and keeping in mind the mapping between different shells and devices.

3 Design Principles of a Modular Framework for Network Troubleshooting

Instead of proposing a new troubleshooting tool or another decision support mechanism, we suggest here a framework¹ capable of *combining* existing (and future) special-purpose tools and reasoning methodologies in a modular fashion. Our concept builds on the observation that operators combine different troubleshooting tools to find out the root cause of a network issue. In the following sections, we go through the main notions that we use to describe such troubleshooting processes and the fundamentals of our framework capable of executing troubleshooting graphs (TSGs).

3.1 Nodes: Wrappers Around Troubleshooting Tools

The first thing we need is an abstraction that incorporates the basic elements of a troubleshooting process. Thus we define our *nodes* as wrappers around troubleshooting tools or smaller, processing functions. Nodes are considered as black boxes that hide their internal operation from the operator (see Fig. 1). Operators have three types of interfaces for communicating with nodes: inputs, on which the nodes execute operations (e.g. a text stream to process or clock ticks), configuration arguments and outputs for relaying information.

The outputs can relay the exact output of the wrapped tool or provide extra processing before generating their results. Nodes are also responsible for providing documentation for these interfaces.

Nodes have three stages of their life cycle: initialization, execution and termination. The initialization stage is where the environment setup of the node is done, including resource allocation and initial configuration. Nodes enter this stage of their life cycle only once. At the execution stage, nodes read the data arriving on their inputs and start the wrapped process or function. Analysis or modification on the wrapped tool’s output is also performed

¹Our initial research and implementation of a subset of current framework functionality is discussed in [24].

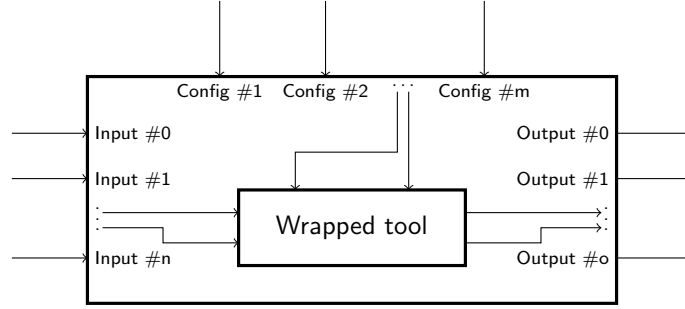


Figure 1: A conceptual node.

here. The node is constantly in this stage when it has been initialized but not yet been terminated. Finally, the node reaches the termination stage when it is being stopped. This stage is responsible for clearing up allocated resources and terminating wrapped processes.

We identified two basic node types: processing nodes and display nodes. The first interprets its inputs and—after processing—it relays the results on its outputs. The second type only interprets input data and is responsible for creating a graphical representation for it, e.g. in the form of a chart or a graph.

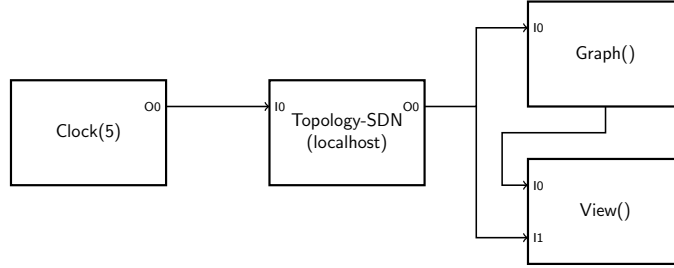
For an example, consider a basic node wrapping the `traceroute` tool. The node outputs the trace (optionally from a remote machine) to a specific host every time we call it. Let’s define first the node’s interfaces:

- Inputs: only a single input is needed for receiving an enabling signal. Each time this input changes, the node should call the `traceroute` command.
- Configuration arguments: we can simply choose to make these the same as the command line arguments of the `traceroute` command. If we want to run `traceroute` on a remote machine, we can add another argument to specify the remote host.
- Outputs: the node should output the results returned by the `traceroute` call on its first output. On the second output, it shows the last hop on the way to the specified target.

In the initialization stage, when requested, the node sets up connection to the remote host where `traceroute` should be run. At execution time, it calls the `traceroute` command and relays its unmodified result to the first output while analyzing it. If the analysis found the last hop, the node sends this information to its second output. In the termination stage, the ongoing processes are stopped and the connection to the remote host is closed, if it had been set up.

3.2 Edges: Accessible Data Transfer

We use *edges* to describe the connections between nodes. Besides specifying the nodes to connect, the main feature required from edges is to provide accessibility for node outputs: information over the edges is observable and modifiable on demand by the operator. When access is provided to edges, the operator can see what is happening in the troubleshooting process on the lowest levels. When historical data is available on the edges, the operator can backtrack how network conditions changed during runtime. Modifiable edges provide the additional benefit of direct operator interaction with nodes, which is helpful for instant testing purposes. Let’s take the example when our troubleshooting scenario is described as a path graph and the first node infrequently outputs data. In this case, the rest of the nodes are also executed rarely and it is quite hard to test whether our graph works as expected.



```

1 Clock(5) -> t :: Topology-SDN(localhost)
2 -> Graph() --> view;
3 t[0] -> [1]view;

```

Figure 2: A simple TSG example (top) and its description (bottom).

When having the option of modifying edge content, we can easily inject test data on the first edge instead of creating a test node that would supply the same data. This method also helps channeling—otherwise unobtainable—data into our system from the network environment.

3.3 The Troubleshooting Graph

Building on the concept of wrapper nodes and accessible edges, a TSG can be created that is able to describe troubleshooting processes as series of tools and transformations. The top part of Fig. 2 depicts a simple TSG that queries, in every five seconds, an SDN controller—accessible on the local machine—for topology information and displays it using a graph visualization node. A special grouping node, the **View**, creates a grouping consisting of a node and an output, in order for these to be displayed together.

Besides the simple concatenation of nodes, creating branches is possible through special purpose *decision* nodes. These nodes are processing nodes able to analyze incoming data and match against a specific criteria set. Such nodes can provide generalized decision making apparatus that can combine results arriving from different nodes and implement arbitrary decision functions to analyze and evaluate them. For example, such a function can validate the output of the `ifconfig` tool and decide whether or not each interface is configured correctly. Another example can be a function that accepts numerical inputs and computes a weighted combination of them, like nodes do in a Bayesian or neural network. We do not consider these decision functions as part of our tool, they can be added to the framework by operators or third parties in a modular fashion. We note that a large collection of decision functions can significantly shorten the creation time of effective branching TSGs.

3.3.1 A Language for Describing TSGs

For a text-based representation of TSGs, we define a simple Click-inspired [25] description language. Such language fits perfectly to our concept as we look at nodes as black boxes that have inputs, outputs and configuration arguments, which the language supports by default. Port-based explicit node linking is also a feature that we make good use of. Nodes can be defined by assigning an instance name to a wrapper node using the `::` operator (see line 1 of Fig. 2). For simplicity, the instance name can be omitted if the node is not referenced in the code later on. Configuration arguments follow the node instance assignment in parentheses (see line 1). Nodes can be linked with edges by linking expressions, using the `->` linking operator. Line 3 of the example shows a port-based explicit linking

of the topology and view nodes. A linking operator always has the list of node outputs on its left side and the list of inputs on its right side, thus multiple edges can be created between two nodes in a single expression. Outputs and inputs are always referred to with their zero-based indices. In case only the 0th input or output appears in a linking expression, one can omit its index, as depicted in line 1 of Fig. 2. In our language—unlike in Click—one can connect outputs to configuration arguments as well, which enables the flexible dynamic configuration of nodes. This uses the same syntax as the output–input linking. To distinguish between inputs and configuration arguments, a hyphen should be prepended to the one-based index of the configuration arguments (see lines 20–21 of Listing 1 for an example). We created a special linking operator, the `-->`, that can be used when connecting the node itself to a View, as used in line 2 of our example². Terminating an expression (a node declaration or a series of linkings) should always be done with a semicolon.

3.4 Execution Framework for TSGs

To bring the TSG concept closer to implementation, we designed an execution framework capable of interpreting TSGs, executing them and providing intelligent navigation among the (possibly many) outputs of the tools used in the TSG.

3.4.1 Interpretation

Since a TSG is only a formal description, the framework provides a parser to interpret the graph. Two methods are available to accomplish this. The first option is building the graph by interpreting its definition written in our description language. Using this method, graphs can be saved and later reused and shared. The parser collects the nodes, their configuration arguments and the links between the nodes. Objects are created for nodes and edges. The framework assigns names to node and edge objects that uniquely identifies them. The second option is creating the graph incrementally at run time using a drag and drop method. This method has the benefit that the operator can monitor the output of the used tools and adjust her methodology to the results she has acquired until that point (which is fully in line with current practice). The parser creates the objects for the nodes and interconnections, and creates a description for these using the language. Run-time modification of the graph is provided by this functionality as well. These two methods provide flexibility while retaining the benefit of being able to store troubleshooting scenarios, situations.

In order to help writing a TSG definition file, the framework provides syntax highlighting to differentiate semantic units. Using the nodes’ self-documentation capabilities, the framework is able to provide on-the-fly node documentation.

3.4.2 Execution

The TSG concept describes how to connect tools to each other but it does not deal with the problem of when and how a node’s life cycle is managed and how a node is notified when its inputs are updated. After interpretation, the framework creates placeholders for nodes which they can then use for displaying graphical data. Once TSG execution is started, the framework is responsible for handling node execution, as depicted in Fig. 3. The framework creates a scheduler that handles a queue for registering node output changes. Each time an output has changed, the framework queries which nodes have that as their input. These destination nodes are then inserted to the end of the queue. Parallel to this, a simple scheduling mechanism is running that always takes the first item from the beginning of the queue and sends a signal to that node to enter into the execution stage. When the call returns, the scheduler moves to the next node in line and so on.

²In this case the **Graph** node does not have any outputs but is responsible for displaying graph visualization on its own.

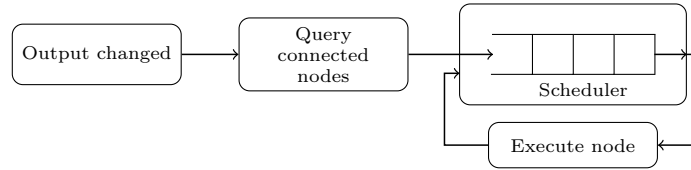


Figure 3: Scheduling of node execution.

3.4.3 Navigation Options

The benefit of handling interconnected troubleshooting tools as a graph is that it creates a natural order in the troubleshooting process. In order to better manage the complex information set contained in the graph, the framework provides different apparatuses to aid observing the execution state and navigating through the graph. A special node class, a **View**, is available that can collect nodes or their outputs and group those to be displayed together. When a name appears without a class assignment, the interpreter assumes that it is a **View**. Layout of such groupings can be explicitly specified or left to the framework to create an automatic one. The framework provides methods for graph traversal by jumping through successive nodes and edges and offers a semantically ordered grouping for nodes and edges contained in the currently displayed TSG. It also provides a visualization method that creates a graphical representation of the TSG.

3.5 Recommendation System and Knowledge Sharing

The framework provides a recommendation system that is able to suggest new nodes for the operator, based on her current setup, by searching for similarities in a TSG repository. Operators can upload their existing TSGs to this repository hereby promoting knowledge sharing.

4 An Everyday Example

For showcasing our framework, we present here an everyday example that (or very similar cases), almost surely, appears in the practice of every operator. We set up a Mininet emulated network to create failures and compare manual troubleshooting with our TSG-based methodology. In our network, hosts are connected to servers via OpenFlow switches and ordinary routers—emulated as hosts having the ability to route between their interfaces. The switches act as simple layer 2 devices, we use only static routes and **iptables** is applied for creating firewall rules. SSH access is set up for every device in the network.

We inject various network related errors to cause the most basic symptom of a network issue: a host on the network is not able to connect to a server. Emulated errors encompass hardware failures—link and port failures—, configuration errors—faulty configuration of hosts, misconfigured routes and firewall rules on routers—and application level errors—misconfigured or unresponsive applications. Our task is to identify and correct the error: first in line with current practice and then, using our framework.

4.1 Current Practice

By applying manual troubleshooting using multiple shells to connect to different devices and running different troubleshooting tools, we made the following observations. We had to repeatedly log into devices when starting a new tool and for every new tool, a new terminal had to be opened. Processes running in currently open terminals could have been terminated in order to minimize the overpopulation of our environment but than we would have

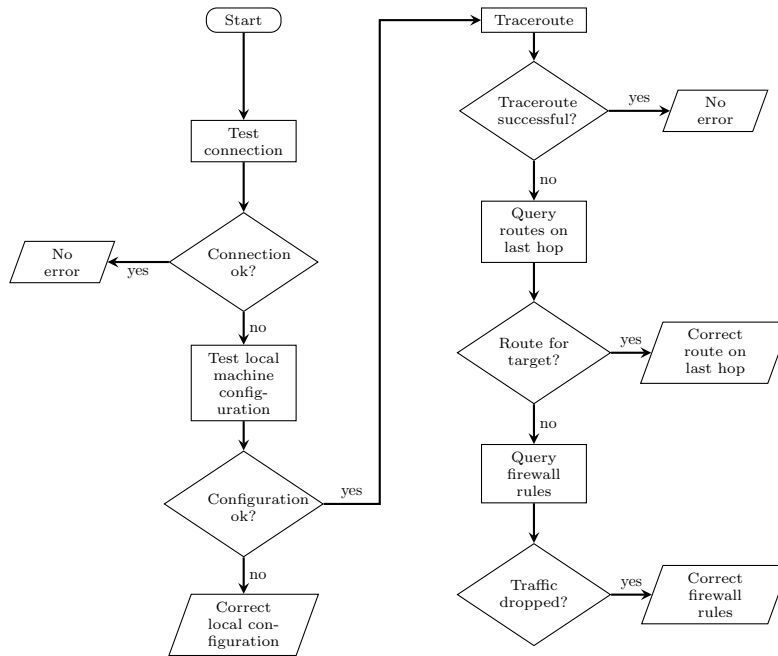


Figure 4: Flow chart for troubleshooting the exemplary scenario.

lost information. We had to analyze the data provided by the tools ourselves and it took a considerable time when we were dealing with great amounts of data. Finally, we noticed that the list of applicable troubleshooting tools could be narrowed down to a minimum and then it did not matter what the root cause was, we ended up using the same tools in the same sequence to find the problem. Thus we came up with a flow chart for this troubleshooting scenario (see Fig. 4). One major problem with the multiple shell approach is that there is no clear way that the process—the steps to be taken—can be recorded and later reused.

To test connectivity we used `ping`. For checking the local host’s configuration we applied the `host`, `ifconfig` and `arp` tools. Forwarding rules were queried by `route`, and firewall rules were retrieved using `iptables`. The steps shown in Fig. 4 could have been grouped together into a single shell script file for automation purposes but we found that method less clear and portable.

4.2 Using TSGs

Now we show that the above process can be easily mapped³ to a TSG, as most of the nodes are already identified by the processes and decisions on the flow chart (Fig. 4). We use extended wrapper nodes for the tools, so—on top of their basic functionality—we assume that `Host`, `Ifconfig` and `Arp` nodes provide means to exclude specified interfaces from their outputs, the `Traceroute` node is able to relay the last hop until which the traffic is traceable and the `Route` and `Iptables` nodes are capable of displaying only those rules that apply to certain IP addresses.

For brevity, Listing 1 shows only⁴ the instructions used for creating a TSG that performs a connection test and if that is unsuccessful, checks the local machine’s network configuration. This exemplary TSG is not complete:

³With an appropriate node repository only the connections need to be correctly specified among nodes.

⁴See the complete TSG in Appendix A.

Listing 1: Formal description of the TSG.

```

1 ping :: Ping(localhost, <address of the server>);
2 ifconfig :: Ifconfig(localhost);
3 arp :: Arp(localhost, nil, -n);
4
5 ping-decision :: Decision(..., string-match,
6                        ..., ttl);
7 ifc-decision :: Decision(...,
8                        ifconfig-check-interfaces,
9                        ..., lo);
10 arp-decision :: Decision(..., (lambda (x)
11                        (> (length x) 0)),
12                        ...);
13
14 ping -> ping-decision;
15 ifconfig -> ifc-decision;
16 arp -> arp-decision;
17
18 ping-decision[1] -> ifconfig;
19 ifc-decision -> Function(ifconfig-get-interfaces,
20                        'input-0)[0, 0]
21 -> [0, -2]arp;
22
23 ping-decision[2] -> ds :: Decision-summary();
24 ifc-decision[2] -> [1]ds;
25 arp-decision[2] -> [2]ds;

```

routes and firewall rules are not checked. Instructions for creating nodes for those tests, however, would be written using the same philosophy shown here. **Decision** nodes in the TSG take the same role as the decisions shown on the flow chart of Fig. 4, while the **Decision-summary** node interprets decisions in the TSG by signaling “error” codes with visual aids for the operator. (Decision nodes are only briefly discussed here, see Section 5.2 for more details.)

In order to perform a connection test between the current host and a server, we use a wrapper node for the **ping** tool in line 1. Lines 2–3 create tests for checking the local host’s configuration using the **ifconfig** and **arp** tools. In order to automatically evaluate these, we defined three Decision nodes. **ping-decision** checks whether the **ping** was successful. **ifc-decision** uses a custom function to validate the interface configuration returned by the **ifconfig** node—the *loopback* interface is ignored by the test for obvious reasons. The **arp-decision** node performs a simple check to test whether there are entries in the Address Resolution Protocol (ARP) cache. These Decision nodes are then connected with their respective wrapper nodes in lines 14–16. In order to check the local host’s configuration only when the connection test was unsuccessful, we need to connect the **ifconfig** node to the negative output of the **ping-decision** node—line 18 implements that. If there are interfaces on the host that are correctly configured, we need to check whether the host can register the layer 2 addresses from its network. We defined a **Function** node in line 19 in order to retrieve the interface names from the output of the **ifc-decision** and fed these to the our **Arp** node. Finally, we defined a **Decision-summary** node in lines 23–25 to display information collected from every Decision node in a summarizing table. A possible output (generated with our prototype) of the Decision-summary node is shown in Fig. 5. The node gives the results of the individual decisions in the TSG as well as an overall evaluation of the current troubleshooting scenario.

This simple example attests that by using TSGs, we can achieve a state of automation where the parametrization of troubleshooting tools adapts to the current network situation and issue. When executing this TSG, the operator

Decision node	Result	Timestamp	Reason
TSG scenario	failed		
ping-decision	failed	@2016-04-13T18:12:15+02:00	every input: timeou
ifc-decision	passed	@2016-04-13T18:12:15+02:00	input-0 (*link:ifco
host-decision	passed	@2016-04-13T18:12:14+02:00	input-0 (*link:host
arp-decision	passed	@2016-04-13T18:12:15+02:00	input-0 (*link:arp:
trace-decision	passed	@2016-04-13T18:12:15+02:00	input-0 (*link:trac
route-decision	passed	@2016-04-13T18:12:15+02:00	input-0 (*link:rout

U:*** *node:ds:Decision-summary* All (1,0) (Epoxide)

Figure 5: Summarizing a troubleshooting scenario.

can recognize failure modes at a glance by looking at the error codes, or can further delve into details by navigating through the outputs of nodes. Using the navigation options provided by the framework, the operator can walk through the troubleshooting process in an orderly fashion. Results are going to be displayed according to the workflow, she laid out when she had designed the flow chart for locating the issue. Even more, the operator has the means to hide irrelevant data by grouping nodes and edges together with Views, or at runtime select only those, she deems relevant to the situation at hand.

By offering proper formalization, TSGs can be reused in similar scenarios. In some cases, the whole graph can be reused with only slight adjustments to the node configuration arguments to adapt to other network conditions. In other cases, subgraphs of the original TSG can be reused for testing different scenarios. Besides re-usability, TSGs can act as a technique to collect troubleshooting know-hows. Once a network problem is uncovered using a TSG, it automatically becomes a guideline for discovering future similar issues. By collecting a library of these, operators can greatly decrease problem solution times and the efficiency of knowledge transfer to new operators can also increase. If TSGs are not only collected within a closed environment—i.e. within a company—but are shared with a greater audience, they can prove to be beneficial for the whole networking community. If a wide TSG library is paired with problem descriptions and solutions, new nodes and test cases can be recommended to an operator, based on previous cases described by TSGs in the library.

5 Prototype

For proofing the concept of our framework, we created a prototype implementation named Epoxide using GNU Emacs as a platform. Emacs is an extensible, customizable text editor, its central concept, the buffer, is responsible for holding file contents, subprocess outputs, providing configuration interfaces, etc. By its extensible nature, Emacs offered a particularly good platform to build Epoxide upon. Emacs supports, via its advanced text manipulation and documentation functions, writing TSG definitions that we store in `.tsg` configuration files. We implemented our execution framework and wrapper nodes using Emacs’s own Lisp dialect, Emacs Lisp. Nodes and their outputs are assigned to Emacs buffers for observability. Node interface and connection information is stored in buffer local variables. We note that while Emacs proved to be a good fit for our notions, the concept described in Section 3 can be implemented on other platforms as well.

5.1 Framework Functions

Opening a `.tsg` file in Emacs automatically loads Epoxide and interprets the TSG stored in it. After assigning node objects and outputs to buffers, nodes can write in these buffers directly during their execution stage. In our

prototype implementation, nodes use other nodes’ output buffers directly as their inputs. Thus TSG edges do not have independent objects.

Emacs’s *ElDoc* package uses the nodes’ self-documentation property to provide hints on node interfaces during TSG definition. When Emacs’s *Auto Complete* package is installed, it can provide intelligent code completion for setting up nodes.

Once a TSG is running, there are two ways to modify edges and node configuration arguments. The first method is runtime TSG reconfiguration, which is provided via an Emacs widget based interface. It allows the addition of TSG edges and modification of static configuration arguments. Modifications are committed to the `.tsg` file as well. The second method a TSGs can be modified is by way of editing the `.tsg` file and reevaluating it. This method has the downside that buffer contents prior to the modification get lost.

Once a TSG is processed, execution is started. The framework stores events in a simple Emacs Lisp list and utilizes the *timerfunctions* package for scheduling node execution. Basic navigability among the created buffers is provided by Emacs key bindings. Epoxide supplies the apparatus to move from one node buffer to its output buffers or to the next node’s buffer (in forward or backward direction). To traverse a TSG, a visualization can be used also, supported by the Emacs *COGRE* package. When the *Graphviz* external software is installed, the displayed graph can be drawn using an automatic layout for better visual clarity. Semantic grouping of the created buffers is provided using the *Ibuffer* package: a dynamic list is displayed that aggregates buffers based on their types and roles in the current context.

The current implementation of Epoxide provides a module for collecting TSG related data and supplying node recommendations. We created the instrumentation for basic case-based reasoning where currently available TSGs are considered as descriptions of previous troubleshooting cases. These TSGs are indexed and their data—together with information from the current TSG—is passed to a recommender. The recommender then can suggest nodes based on these pieces of data. We created an interface in the framework to which recommenders, developed by third parties, can connect as well. By now, we have implemented the most basic recommender that suggests the most popular nodes and displays them using Emacs’s *Ido* package. Most popular nodes are computed by counting every node in all previously written TSGs and ordering them in the descending order of their cumulative count. Nodes that are already used in the current TSG, are excluded from the suggestion list.

5.2 Branching Nodes

When creating the apparatus that enables conditional branching in Epoxide, the most basic expectation was to (i) provide functionality to analyze and evaluate the output of any node and, based on the result, (ii) create branches in a TSG, the same way a decision would in a flow chart. Additionally, we wanted to have the ability to (iii) select among different inputs or to use a combination of them. This criterion was inspired by how nodes work in a Bayesian network: they receive the results of lower level nodes and calculate their outputs based on that. To satisfy these criteria, we implemented a single *Decision node*. Since nodes can be added to Epoxide in a modular fashion, this is only one possible implementation that satisfies our initial criteria. Operators are free to add their own version. A *Decision-summary node* was developed in conjunction, for summarizing the outputs of such nodes.

Fig. 6 shows the schematics of our Decision node. The node can attach to any wrapper node⁵ and incoming data is first verified (as per (i)): it is determined whether it complies with certain criteria. The node can use any function for verification that has at least one argument (the input) and can return false, when verification failed, or any string otherwise. When verification of the inputs is finished, further processing can be done using a second stage function, in accordance with (iii). An operator can use a function to select an input with, for example, the *or* function: the first of the inputs that passed verification is going to be the result of this stage. Other functions can implement more complex processing, like in the aforementioned Bayesian network example where the result would

⁵The node is also prepared to handle the asynchronism and the different output formats of the wrapper nodes.

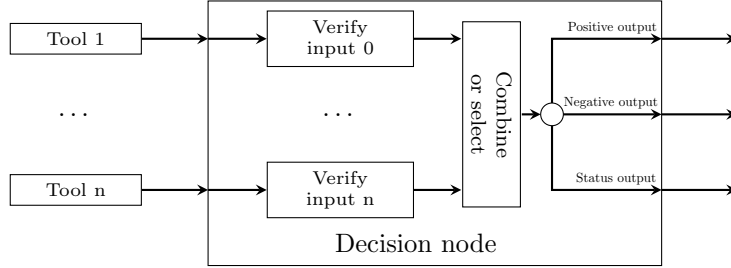


Figure 6: Schematics of a Decision node.

be a probability value. Such second stage functions should return a string in case of a positive decision or false otherwise. We defined two node outputs to relay these return values (as per (ii)). The first displays information in the positive case, the second in the negative. We implemented an additional output for interoperating with Decision-summary nodes.

5.3 Implementing Nodes

The modular architecture of Epoxide makes it possible for anyone to implement new nodes⁶. A node developer—who implements a node—has to create separate node definition files that contain Emacs Lisp code to provide self-documentation for nodes and implement the three life cycle stages via functions. For proper interaction with nodes, the node definition files and these functions should comply with a fixed naming scheme. Node self-documentation functions should provide information about node interfaces and could also implement validation functions to check the compliance of arguments with certain criteria. The functions implementing the separate node life cycle stages are called by the framework on specific occasions. The initialization function is evoked after the buffers belonging to the node are set up. The execution function is called every time a change occurs on any of the node’s inputs. Finally, the termination function is invoked before the framework closes the buffers associated with the node. These functions can draw on common node functions provided by the framework. Epoxide implements functions for setting up node buffers with basic data, reading inputs and configuration arguments, writing outputs and handling remote access using Emacs’s *Transparent Remote Access, Multiple Protocols (TRAMP)* package.

6 Case studies

The basic example shown in Section 4 demonstrated integration of troubleshooting tools originally developed for traditional networks. Although these can also be used to troubleshoot SDNs, tools created specifically for these networks can leverage the benefits of centralized network control.

6.1 Troubleshooting in SDNs

We implemented nodes to interact with POX, Floodlight and OpenDaylight controllers and Open vSwitch (OVS) switches. These nodes are able to collect datapath identifiers (DPIDs), flow statistics and topology information. Additionally, we created processing nodes for filtering flow statistics on a flow space⁷ basis and for visualizing topology information. In order to aid controller debugging, we wrapped the GDB tool as well. To support other

⁶The list of nodes currently implemented in Epoxide is shown in Appendix B.

⁷Those flow table entries make up a flow space that match given source and destination point pairs.

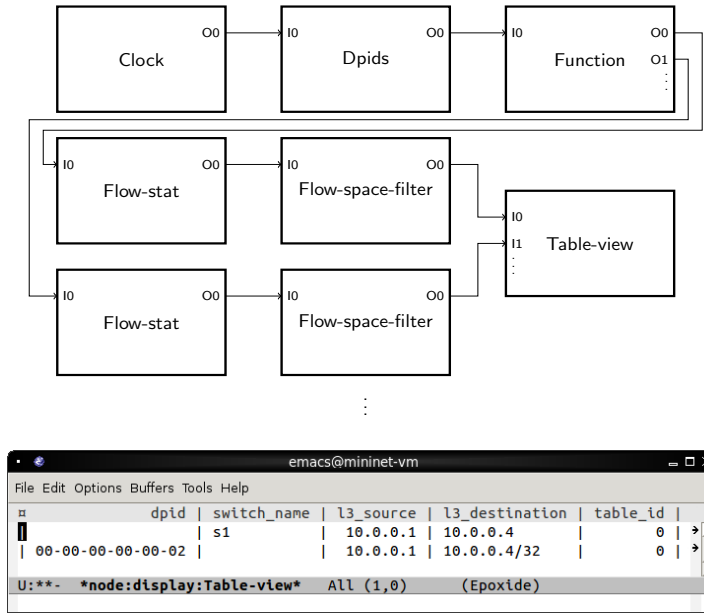


Figure 7: An SDN example: querying flow statistics with a TSG (top) and displaying them in a table using a **Table-view** node (bottom).

controller platforms, a general node for accessing REST APIs was implemented, as well as other nodes that are able to filter specific parts of JSON-formatted textual data.

To illustrate an SDN use-case, imagine the following scenario: a network consisting of SDN switches is given and an operator wants to monitor a certain traffic flow in the network. She can set up a TSG to solve this problem by first defining a node querying DPIDs from the controller on a timely basis, as depicted in the top part of Fig. 7. A **Function** node can then separate the list of available DPIDs and pass those to further nodes that query flow statistics from their assigned switches. With the addition of flow space filtering nodes, the operator can select only the flow under scrutiny and display the results in a table format, as depicted by the bottom part of Fig. 7. She can investigate further by adding nodes that query flow statistics from the controller and from the switches themselves, and compare the results with **Decision** nodes: this could reveal synchronization issues between the devices or misdirection of the traffic caused by a software failure. By connecting **Gdb** nodes with the **Decision** nodes, the operator can remotely connect to software switches or to the controller for an in-depth analyzation of the problem.⁸

6.2 Troubleshooting in Service Function Chaining

One of the main goals of the UNIFY project⁹ was to design a Service Function Chaining control plane architecture and implement a proof-of-concept prototype. Additionally, the project also introduces the concept of Service-Provider DevOps to combine the developer and operational workflows in carrier grade environments. DevOps results in faster deployment cycle of novel networking services. Instead of designing complex services as a whole, these services are assembled from atomic *network functions*. However, fast deployment cycles require faster testing phases

⁸A live action demo on a similar, albeit simpler, case can be watched at [26]: <https://www.youtube.com/watch?v=HsiGFR0QirE>

⁹<https://www.fp7-unify.eu>

and troubleshooting in the operational environment. Even when a new service is created by re-using components of previous ones, it is still going to be different enough from earlier scenarios to implicate new troubleshooting challenges. In these cases, previous knowledge is not always directly applicable. By providing an integrated platform for running troubleshooting tools and an apparatus to automatize their execution, our tool makes the formation of new troubleshooting scenarios easier, thus enabling quicker service deployment.

Epoxide has a central role in the multipurpose demonstrator showcasing major results of the project [27]. Using its dedicated and general purpose wrapper nodes, it orchestrates multiple components in a semi-automatic troubleshooting scenario. Epoxide needs to reveal a configuration error resulting in erroneous imbalance of the traffic loads of OpenFlow switches instantiated as network functions. The novel components, which Epoxide interacts with, include a flexible messaging bus (Double Decker), a tool that calculates aggregated performance metrics derived from data queried from hierarchical time series databases [Recursive Query Engine (RQE)], and a test packet generator for pinpointing errors in OpenFlow switches (AutoTPG).¹⁰

Fig. 8 highlights the main system components and the sequence of interactions. First, a monitoring component detects the resource imbalance that automatically triggers the execution of the troubleshooting process in Epoxide that executes a TSG tailored for this scenario (step 1). In step 2, Epoxide asks the Recursive Query Engine to narrow down the location of the error to a subset of OpenFlow switches. After querying historical measurement data (steps 3–4), RQE returns a list of switches to Epoxide (step 5). In step 6, Epoxide starts and configures AutoTPG and asks it to test candidate switches one by one. AutoTPG tests correctness of the switches in steps 7–8 and returns the result to Epoxide in step 9. Finally, Epoxide queries the flow-entries of the erroneous switch (step 10), and presents those in tabular form to the user to help further investigating the problem manually.

The hundred-line long TSG of the demo contains a wrapper node for the messaging bus, but communicates with the other tools via **Rest-api** nodes. REST API calls request JSON-formatted inputs and emit JSON-formatted outputs. While most of the nodes discussed so far assumed unstructured plain text inputs and outputs, in this case we needed **Json-filter** nodes to prepare the outputs for post-processing, and **Format** nodes to assemble JSON-formatted inputs for the API calls. This example shows that there is not too much difference between a TSG processing structured JSON data and a TSG processing unstructured data line-by-line: similar filter nodes are necessary in both cases. The TSG also exemplifies the usage of a **Tee** node and the **Command** node. Similarly to its Unix counterpart, **Tee** copies its input text to its output link and saves the text in a file. **Command** is used for running an *ssh* command to modify the configuration of a remote network function.

Epoxide exhibits properties that make it an ideal testing and troubleshooting tool for Service Function Chaining. First, the TSG language is an enabler of fast hypotheses testing and small feedback cycles because it allows connecting existing special purpose troubleshooting tools at an abstract level. Second, the test generation process can be further shortened by simply re-using parts of existing **.tsg** definitions. Third, complex decision logics can be based on service-specific monitoring and troubleshooting tools by writing simple wrapper nodes around these tools.

7 Conclusion and Future Work

While our modular framework proposed here is capable of flexibly combining various troubleshooting tools for tracking down networking issues, and the TSG concept enables the accumulation and sharing of troubleshooting related knowledge, the current prototype implementation should be extended in many aspects. Of course, a large library of wrapper nodes should be added to incorporate more and more troubleshooting tools. Besides this natural option, we outline here some future directions, the framework could benefit from.

¹⁰Detailed description of these tools and the demonstrator can be found in [27].

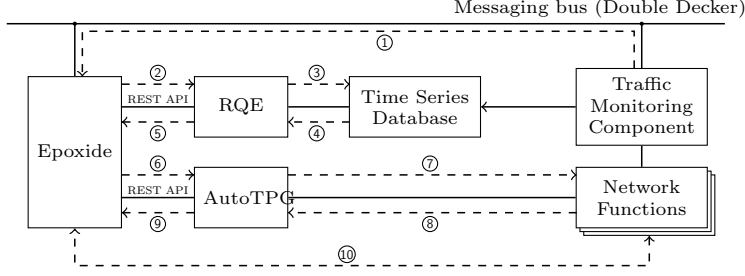


Figure 8: Simplified view of the UNIFY demo architecture and sequence of interactions.

Although present implementation supports the addition of new node recommenders, the currently implemented one provides only a basic functionality. We consider this a good basis to implement better recommenders. The first in line is a recommender that analyzes the current TSG and—by finding the most similar TSG from a collection—suggests new nodes based on that. This concept could be further improved by tagging the TSGs as to what sort of problems they were used for finding out and taking into account the tags during recommendation. Suggestions could be made more relevant if the environment of the nodes were to be taken into consideration as well, and node configuration arguments could be suggested also.

To support our case-based reasoning approach, we imagine a repository of TSGs accessible by anyone for viewing old cases and uploading new ones. At the same time such a repository is set up, recommenders could be moved to the cloud as well, resulting in a more efficient computation of suggestions on a much greater collection of TSGs compared to a locally computed model.

Since case-based reasoning is not the cutting edge of decision support systems, we can use the TSG concept as a failure detection graph to be used as a Bayesian network, and make node suggestions based on that. Finding an applicable failure propagation model is key to this approach. We believe, we can take this concept even further by using a Bayesian network based method and combining it with active diagnostics—where new network tests are selected automatically depending on current results—for TSGs to be built with little operator intervention or totally unsupervised.

Acknowledgment

The research leading to these results was partly supported by Ericsson and has received funding from the European Union Seventh Framework Programme under grant agreement N^o 619609.

A Complete TSG Belonging to the Everyday Example

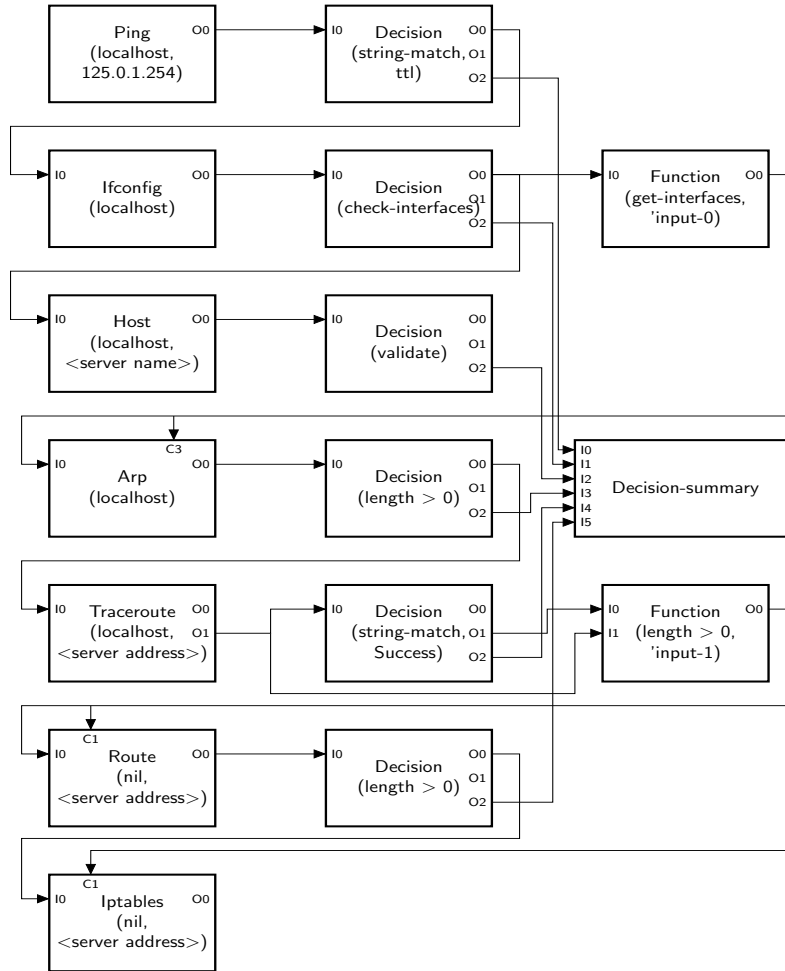


Figure 9: TSG for the example.

B List of Nodes Currently Implemented in Epoxide

Arp:	queries the ARP cache by evoking the arp command.
Clock:	provides a signal on a timely basis.
Command:	wraps any shell commands. These are run as Emacs subprocesses.
Decision:	provides conditional branching.
Decision-summary:	summarizes the results of connected Decision nodes.
Doubledecker:	connects to a DoubleDecker broker and is able to receive and send messages to a partner or to a topic.
Dpids-<controller type>:	collects DPID (Datapath ID) and switch name information from an OpenFlow controller.
Emacs-buffer:	provides options to channel information from any Emacs buffer to Epoxide.
Escape:	queries NFFG topology information from Escape.
Filter:	marks incoming text that matches a given regular expression.
Flow-space-filter:	provides support to select only that flow space that is wished to be seen.
Flow-stat-<controller type>:	provides functionality to query an OpenFlow controller for flow statistics of a specific switch given with its DPID.
Format:	collects text from its inputs and reformats them.
Function:	wraps any Emacs Lisp function or nameless function.
Gdb:	attaches a wrapped GDB to a running process.
Graph:	provides graph visualization support.
Host:	performs DNS lookup using the host shell-call.
Ifconfig:	wraps the ifconfig shell command.
Iperf:	wraps around the iperf command.
Json-filter:	finds the values belonging to a given key in a JSON expression.
Ping:	wraps the ping command.
Rest-api:	performs REST API calls.
Route:	wraps the route command.
Table-view:	displays data received on its inputs in a table form.
Tee:	wraps around the Unix tee command: while forwarding incoming data it also saves it to a given file.
Topology-<controller type>:	queries topology information from an OpenFlow controller.
Traceroute:	wraps the traceroute process.

Table 1: List of implemented Epoxide nodes.

References

- [1] J. D. Sloan, *Network Troubleshooting Tools*. O'Reilly, 8 2001.
- [2] C. Abrahamson, M. Blodgett, A. Kunen, N. Mueller, and D. Parter, "Splat: A Network Switch/Port Configuration Management Tool," in *Proceedings of the 17th Conference on Systems Administration (LISA 2003)*, 2003.
- [3] P. Kazemian, M. Chan, H. Zeng, G. Varghese, N. McKeown, and S. Whyte, "Real Time Network Policy Checking Using Header Space Analysis," in *NSDI*, 2013, pp. 99–111.
- [4] P. Kazemian, G. Varghese, and N. McKeown, "Header Space Analysis: Static Checking for Networks," in *NSDI*, 2012, pp. 113–126.
- [5] H. Zeng, P. Kazemian, G. Varghese, and N. McKeown, "Automatic test packet generation," in *Proceedings of the 8th international conference on Emerging networking experiments and technologies*. ACM, 2012, pp. 241–252.
- [6] H. Mai, A. Khurshid, R. Agarwal, M. Caesar, P. B. Godfrey, and S. T. King, "Debugging the Data Plane with Anteater," in *Proceedings of the ACM SIGCOMM 2011 Conference*, ser. SIGCOMM '11. New York, NY, USA: ACM, 2011, pp. 290–301. [Online]. Available: <http://doi.acm.org/10.1145/2018436.2018470>

- [7] A. Wundsam, D. Levin, S. Seetharaman, A. Feldmann *et al.*, “OFRewind: Enabling Record and Replay Troubleshooting for Networks,” in *USENIX Annual Technical Conference*, 2011.
- [8] N. Handigol, B. Heller, V. Jeyakumar, D. Mazieres, and N. McKeown, “I know what your packet did last hop: Using packet histories to troubleshoot networks,” in *Proc. USENIX NSDI*, 2014.
- [9] A. Khurshid, W. Zhou, M. Caesar, and P. Godfrey, “Veriflow: verifying network-wide invariants in real time,” *ACM SIGCOMM Computer Communication Review*, vol. 42, no. 4, pp. 467–472, 2012.
- [10] M. Canini, D. Venzano, P. Peresini, D. Kostic, J. Rexford *et al.*, “A NICE Way to Test OpenFlow Applications,” in *NSDI*, vol. 12, 2012, pp. 127–140.
- [11] M. Kuzniar, P. Peresini, M. Canini, D. Venzano, and D. Kostic, “A soft way for openflow switch interoperability testing,” in *Proceedings of the 8th international conference on Emerging networking experiments and technologies*. ACM, 2012, pp. 265–276.
- [12] P. Porras, S. Shin, V. Yegneswaran, M. Fong, M. Tyson, and G. Gu, “A security enforcement kernel for openflow networks,” in *Proceedings of the first workshop on Hot topics in software defined networks*. ACM, 2012, pp. 121–126.
- [13] M. Kuzniar, M. Canini, and D. Kostic, “OFTEN testing OpenFlow networks,” in *Software Defined Networking (EWSN), 2012 European Workshop on*. IEEE, 2012, pp. 54–60.
- [14] R. Mathonet, H. V. Cotthem, and L. Vanryckeghem, “DANTES An Expert System for Real-Time Network Troubleshooting,” in *Proceedings of the 10th International Joint Conference on Artificial Intelligence*, 1987, pp. 527–530.
- [15] B. L. Hitson, “Knowledge-Based Monitoring and Control: An Approach to Understanding the Behavior TCP/IP Network Protocols,” in *Symposium proceedings on Communications architectures and protocols*, vol. 18, 1988.
- [16] L. Lewis, “A Case-Based Reasoning Approach to the Management of Faults in Communications Networks,” in *Twelfth Annual Joint Conference of the IEEE Computer and Communications Societies*, 1993.
- [17] G. Jakobson and M. Weissman, “Real-time telecommunication network management: extending event correlation with temporal constraints,” in *Proceedings of the fourth international symposium on Integrated network management IV*, 1995, pp. 290–301.
- [18] G. Reali and L. Monacelli, “Definition and performance evaluation of a fault localization technique for an NGN IMS network,” in *IEEE Transactions on Network and Service Management*, 2009, p. 6(2):122136.
- [19] J. Lu, C. Dousson, B. Radier, and F. Krief, “Towards an Autonomic Network Architecture for Self-healing in Telecommunications Networks,” in *Mechanisms for Autonomous Management of Networks and Services*, vol. 6155, 2010, pp. 110–113.
- [20] —, “A Self-diagnosis Algorithm Based on Causal Graphs,” in *The Seventh International Conference on Autonomic and Autonomous Systems*, 2011.
- [21] Bayesian networks without tears: making Bayesian networks more accessible to the probabilistically unsophisticated, “Eugene charniak,” in *AI Magazine*, vol. 12, 1991, pp. 50–63.
- [22] R. Khanafer, “Automated diagnosis for UMTS networks using Bayesian network approach,” in *IEEE Transactions on Vehicular Technology*, 2008, p. 57:24512461.
- [23] H. Zeng, P. Kazemian, G. Varghese, and N. McKeown, “A Survey on Network Troubleshooting,” 2014.
- [24] I. Pelle, T. Lévai, F. Németh, and A. Gulyás, “One tool to rule them all: A modular troubleshooting framework for SDN (and other) networks,” in *Proceedings of the 1st ACM SIGCOMM Symposium on Software Defined Networking Research*, 2015.
- [25] E. Kohler, R. Morris, B. Chen, J. Jannotti, and M. F. Kaashoek, “The Click modular router,” *ACM Transactions on Computer Systems (TOCS)*, vol. 18, no. 3, pp. 263–297, 2000.
- [26] T. Lévai, I. Pelle, F. Németh, and A. Gulyás, “Epoxide: A modular prototype for sdn troubleshooting,” in *Proceedings of the 2015 ACM Conference on Special Interest Group on Data Communication*, 2015, pp. 359–360.
- [27] G. Marchetto and R. Sisto (editors), “Deliverable D4.3: Updated concept and evaluation results for SP-DevOps,” to appear, UNIFY Project, Tech. Rep., 2016.